



ENTERED

PCT09

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/830,706B

DATE: 04/22/2002
TIME: 15:01:38

Input Set : A:\55865sequence.txt
Output Set: N:\CRF3\04222002\I830706B.raw

4 <110> APPLICANT: TOJI, SHINGO
5 YANO, MINORU
6 TAMAI, KATSUYUKI
8 <120> TITLE OF INVENTION: THIOREDOXIN REDUCTASE II
10 <130> FILE REFERENCE: 55865-71965
12 <140> CURRENT APPLICATION NUMBER: 09/830,706B
C--13 <141> CURRENT FILING DATE: 2002-01-29
15 <150> PRIOR APPLICATION NUMBER: PCT/JP99/05983
16 <151> PRIOR FILING DATE: 1999-10-28
18 <150> PRIOR APPLICATION NUMBER: JP 1998-310422
19 <151> PRIOR FILING DATE: 1998-10-30
21 <160> NUMBER OF SEQ ID NOS: 38
23 <170> SOFTWARE: PatentIn Ver. 2.1
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 1959
27 <212> TYPE: DNA
28 <213> ORGANISM: Homo sapiens
30 <220> FEATURE:
31 <221> NAME/KEY: CDS
32 <222> LOCATION: (10)..(1572)
34 <220> FEATURE:
35 <221> NAME/KEY: misc_feature
36 <222> LOCATION: (1567)..(1569)
37 <223> OTHER INFORMATION: "tga" is translated to selenocysteine
39 <400> SEQUENCE: 1
40 atggcgga atg gcg gtg gcg ctg cgg gga tta gga ggg cgc ttc cgg tgg 51
41 Met Ala Val Ala Leu Arg Gly Leu Gly Gly Arg Phe Arg Trp
42 1 5 10
44 cgg acg cag gcc gtg gcg ggc ggg gtg cgg ggc gcg gcg cgg ggc gca 99
45 Arg Thr Gln Ala Val Ala Gly Gly Val Arg Gly Ala Ala Arg Gly Ala
46 15 20 25 30
48 gca gca ggt cag cgg gac tat gat ctc ctg gtg gtc ggc ggg gga tct 147
49 Ala Ala Gly Gln Arg Asp Tyr Asp Leu Leu Val Val Gly Gly Gly Ser
50 35 40 45
52 ggt ggc ctg gct tgt gcc aag gag gcc gcc cag ctg gga agg aag gtg 195
53 Gly Gly Leu Ala Cys Ala Lys Glu Ala Ala Gln Leu Gly Arg Lys Val
54 50 55 60
56 gcc gtg gtg gac tac gtg gaa cct tct ccc caa ggc acc cgg tgg ggc 243
57 Ala Val Val Asp Tyr Val Glu Pro Ser Pro Gln Gly Thr Arg Trp Gly
58 65 70 75
60 ctc ggc ggc acc tgc gtc aac gtg ggc tgc atc ccc aag aag ctg atg 291
61 Leu Gly Gly Thr Cys Val Asn Val Gly Cys Ile Pro Lys Lys Leu Met
62 80 85 90

RAW SEQUENCE LISTING

DATE: 04/22/2002

PATENT APPLICATION: US/09/830,706B

TIME: 15:01:38

Input Set : A:\55865sequence.txt

Output Set: N:\CRF3\04222002\I830706B.raw

```

64 cac cag gcg gca ctg ctg gga ggc ctg atc caa gat gcc ccc aac tat 339
65 His Gln Ala Ala Leu Leu Gly Gly Leu Ile Gln Asp Ala Pro Asn Tyr
66 95 100 105 110
68 ggc tgg gag gtg gcc cag ccc gtg ccg cat gac tgg agg aag atg gca 387
69 Gly Trp Glu Val Ala Gln Pro Val Pro His Asp Trp Arg Lys Met Ala
70 115 120 125
72 gaa gct gtt caa aat cac gtg aaa tcc ttg aac tgg ggc cac cgt gtc 435
73 Glu Ala Val Gln Asn His Val Lys Ser Leu Asn Trp Gly His Arg Val
74 130 135 140
76 cag ctt cag gac aga aaa gtc aag tac ttt aac atc aaa gcc agc ttt 483
77 Gln Leu Gln Asp Arg Lys Val Lys Tyr Phe Asn Ile Lys Ala Ser Phe
78 145 150 155
80 gtt gac gag cac acg gtt tgc ggc gtt gcc aaa ggt ggg aaa gag att 531
81 Val Asp Glu His Thr Val Cys Gly Val Ala Lys Gly Gly Lys Glu Ile
82 160 165 170
84 ctg ctg tca gcc gat cac atc atc att gct act gga ggg cgg ccg aga 579
85 Leu Leu Ser Ala Asp His Ile Ile Ile Ala Thr Gly Gly Arg Pro Arg
86 175 180 185 190
88 tac ccc acg cac atc gaa ggt gcc ttg gaa tat gga atc aca agt gat 627
89 Tyr Pro Thr His Ile Glu Gly Ala Leu Glu Tyr Gly Ile Thr Ser Asp
90 195 200 205
92 gac atc ttc tgg ctg aag gaa tcc cct gga aaa acg ttg gtg gtc ggg 675
93 Asp Ile Phe Trp Leu Lys Glu Ser Pro Gly Lys Thr Leu Val Val Gly
94 210 215 220
96 gcc agc tat gtg gcc ctg gag tgt gct ggc ttc ctc acc ggg att ggg 723
97 Ala Ser Tyr Val Ala Leu Glu Cys Ala Gly Phe Leu Thr Gly Ile Gly
98 225 230 235
100 ctg gac acc acc atc atg atg cgc agc atc ccc ctc cgc ggc ttc gac 771
101 Leu Asp Thr Thr Ile Met Met Arg Ser Ile Pro Leu Arg Gly Phe Asp
102 240 245 250
104 cag caa atg tcc tcc atg gtc ata gag cac atg gca tct cat ggc acc 819
105 Gln Gln Met Ser Ser Met Val Ile Glu His Met Ala Ser His Gly Thr
106 255 260 265 270
108 cgg ttc ctg agg ggc tgt gcc ccc tcg cgg gtc agg agg ctc cct gat 867
109 Arg Phe Leu Arg Gly Cys Ala Pro Ser Arg Val Arg Arg Leu Pro Asp
110 275 280 285
112 ggc cag ctg cag gtc acc tgg gag gac agc acc acc ggc aag gag gac 915
113 Gly Gln Leu Gln Val Thr Trp Glu Asp Ser Thr Thr Gly Lys Glu Asp
114 290 295 300
116 acg ggc acc ttt gac acc gtc ctg tgg gcc ata ggt cga gtc cca gac 963
117 Thr Gly Thr Phe Asp Thr Val Leu Trp Ala Ile Gly Arg Val Pro Asp
118 305 310 315
120 acc aga agt ctg aat ttg gag aag gct ggg gta gat act agc ccc gac 1011
121 Thr Arg Ser Leu Asn Leu Glu Lys Ala Gly Val Asp Thr Ser Pro Asp
122 320 325 330
124 act cag aag atc ctg gtg gac tcc cgg gaa gcc acc tct gtg ccc cac 1059
125 Thr Gln Lys Ile Leu Val Asp Ser Arg Glu Ala Thr Ser Val Pro His
126 335 340 345 350
128 atc tac gcc att ggt gac gtg gtg gag ggg cgg cct gag ctg aca ccc 1107

```

RAW SEQUENCE LISTING

DATE: 04/22/2002

PATENT APPLICATION: US/09/830,706B

TIME: 15:01:38

Input Set : A:\55865sequence.txt

Output Set: N:\CRF3\04222002\I830706B.raw

129 Ile Tyr Ala Ile Gly Asp Val Val Glu Gly Arg Pro Glu Leu Thr Pro
130 355 360 365
132 aca gcg atc atg gcc ggg agg ctc ctg gtg cag cgg ctc ttc ggc ggg 1155
133 Thr Ala Ile Met Ala Gly Arg Leu Leu Val Gln Arg Leu Phe Gly Gly
134 370 375 380
136 tcc tca gat ctg atg gac tac gac aat gtt ccc acg acc gtc ttc acc 1203
137 Ser Ser Asp Leu Met Asp Tyr Asp Asn Val Pro Thr Thr Val Phe Thr
138 385 390 395
140 cca ctg gag tat ggc tgt gtg ggg ctg tcc gag gag gag gca gtg gct 1251
141 Pro Leu Glu Tyr Gly Cys Val Gly Leu Ser Glu Glu Glu Ala Val Ala
142 400 405 410
144 cgc cac ggg cag gag cat gtt gag gtc tat cac gcc cat tat aaa cca 1299
145 Arg His Gly Gln Glu His Val Glu Val Tyr His Ala His Tyr Lys Pro
146 415 420 425 430
148 ctg gag ttc acg gtg gct gga cga gat gca tcc cag tgt tat gta aag 1347
149 Leu Glu Phe Thr Val Ala Gly Arg Asp Ala Ser Gln Cys Tyr Val Lys
150 435 440 445
152 atg gtg tgc ctg agg gag ccc cca cag ctg gtg ctg ggc ctg cat ttc 1395
153 Met Val Cys Leu Arg Glu Pro Pro Gln Leu Val Leu Gly Leu His Phe
154 450 455 460
156 ctt ggc ccc aac gca ggc gaa gtt act caa gga ttt gct ctg ggg atc 1443
157 Leu Gly Pro Asn Ala Gly Glu Val Thr Gln Gly Phe Ala Leu Gly Ile
158 465 470 475
160 aag tgt ggg gct tcc tat gcg cag gtg atg cgg acc gtg ggt atc cat 1491
161 Lys Cys Gly Ala Ser Tyr Ala Gln Val Met Arg Thr Val Gly Ile His
162 480 485 490
164 ccc aca tgc tct gag gag gta gtc aag ctg cgc atc tcc aag cgc tca 1539
165 Pro Thr Cys Ser Glu Glu Val Val Lys Leu Arg Ile Ser Lys Arg Ser
166 495 500 505 510
168 ggc ctg gac ccc acg gtg aca ggc tgc tga ggg taagcgccat ccctgcaggc 1592
W--> 169 Gly Leu Asp Pro Thr Val Thr Gly Cys Xaa Gly
170 515 520
172 cagggcacac ggtgcgccc cgcgcagctc ctgcggagcc agaccagga tggctgcagg 1652
174 ccaggttttg ggggcctcaa ccctctctg gagcgctgt gagatggtca gcgtggagcg 1712
176 caagtgcctg acgggtggcc cgtgtgcccc acagggatgg ctccaggggac tgtccacctc 1772
178 acccctgcac ctttcagcct ttgcgcgcgg gcaccccccc caggctcctg gtgccggatg 1832
180 atgacgacct ggggtgaaac ctaccctgtg ggcacccatg tccgagcccc ctggcatttc 1892
182 tgcaatgcaa ataaagaggg tactttttct gaagtgtgta aaaaaaaaaa aaaaaaaaaa 1952
184 aaaaaaa 1959
187 <210> SEQ ID NO: 2
188 <211> LENGTH: 521
189 <212> TYPE: PRT
190 <213> ORGANISM: Homo sapiens
192 <220> FEATURE:
193 <221> NAME/KEY: SITE
194 <222> LOCATION: (520)
195 <223> OTHER INFORMATION: Selenocysteine
197 <400> SEQUENCE: 2
198 Met Ala Val Ala Leu Arg Gly Leu Gly Gly Arg Phe Arg Trp Arg Thr

RAW SEQUENCE LISTING

DATE: 04/22/2002

PATENT APPLICATION: US/09/830,706B

TIME: 15:01:38

Input Set : A:\55865sequence.txt

Output Set: N:\CRF3\04222002\I830706B.raw

```

199      1          5          10          15
201 Gln Ala Val Ala Gly Gly Val Arg Gly Ala Ala Arg Gly Ala Ala Ala
202                20          25          30
204 Gly Gln Arg Asp Tyr Asp Leu Leu Val Val Gly Gly Gly Ser Gly Gly
205                35          40          45
207 Leu Ala Cys Ala Lys Glu Ala Ala Gln Leu Gly Arg Lys Val Ala Val
208                50          55          60
210 Val Asp Tyr Val Glu Pro Ser Pro Gln Gly Thr Arg Trp Gly Leu Gly
211      65          70          75          80
213 Gly Thr Cys Val Asn Val Gly Cys Ile Pro Lys Lys Leu Met His Gln
214                85          90          95
216 Ala Ala Leu Leu Gly Gly Leu Ile Gln Asp Ala Pro Asn Tyr Gly Trp
217                100         105         110
219 Glu Val Ala Gln Pro Val Pro His Asp Trp Arg Lys Met Ala Glu Ala
220                115         120         125
222 Val Gln Asn His Val Lys Ser Leu Asn Trp Gly His Arg Val Gln Leu
223                130         135         140
225 Gln Asp Arg Lys Val Lys Tyr Phe Asn Ile Lys Ala Ser Phe Val Asp
226      145         150         155         160
228 Glu His Thr Val Cys Gly Val Ala Lys Gly Gly Lys Glu Ile Leu Leu
229                165         170         175
231 Ser Ala Asp His Ile Ile Ile Ala Thr Gly Gly Arg Pro Arg Tyr Pro
232                180         185         190
234 Thr His Ile Glu Gly Ala Leu Glu Tyr Gly Ile Thr Ser Asp Asp Ile
235                195         200         205
237 Phe Trp Leu Lys Glu Ser Pro Gly Lys Thr Leu Val Val Gly Ala Ser
238                210         215         220
240 Tyr Val Ala Leu Glu Cys Ala Gly Phe Leu Thr Gly Ile Gly Leu Asp
241      225         230         235         240
243 Thr Thr Ile Met Met Arg Ser Ile Pro Leu Arg Gly Phe Asp Gln Gln
244                245         250         255
246 Met Ser Ser Met Val Ile Glu His Met Ala Ser His Gly Thr Arg Phe
247                260         265         270
249 Leu Arg Gly Cys Ala Pro Ser Arg Val Arg Arg Leu Pro Asp Gly Gln
250                275         280         285
252 Leu Gln Val Thr Trp Glu Asp Ser Thr Thr Gly Lys Glu Asp Thr Gly
253                290         295         300
255 Thr Phe Asp Thr Val Leu Trp Ala Ile Gly Arg Val Pro Asp Thr Arg
256      305         310         315         320
258 Ser Leu Asn Leu Glu Lys Ala Gly Val Asp Thr Ser Pro Asp Thr Gln
259                325         330         335
261 Lys Ile Leu Val Asp Ser Arg Glu Ala Thr Ser Val Pro His Ile Tyr
262                340         345         350
264 Ala Ile Gly Asp Val Val Glu Gly Arg Pro Glu Leu Thr Pro Thr Ala
265                355         360         365
267 Ile Met Ala Gly Arg Leu Leu Val Gln Arg Leu Phe Gly Gly Ser Ser
268                370         375         380
270 Asp Leu Met Asp Tyr Asp Asn Val Pro Thr Thr Val Phe Thr Pro Leu
271      385         390         395         400

```

RAW SEQUENCE LISTING

DATE: 04/22/2002

PATENT APPLICATION: US/09/830,706B

TIME: 15:01:38

Input Set : A:\55865sequence.txt

Output Set: N:\CRF3\04222002\I830706B.raw

273 Glu Tyr Gly Cys Val Gly Leu Ser Glu Glu Glu Ala Val Ala Arg His
 274 405 410 415
 276 Gly Gln Glu His Val Glu Val Tyr His Ala His Tyr Lys Pro Leu Glu
 277 420 425 430
 279 Phe Thr Val Ala Gly Arg Asp Ala Ser Gln Cys Tyr Val Lys Met Val
 280 435 440 445
 282 Cys Leu Arg Glu Pro Pro Gln Leu Val Leu Gly Leu His Phe Leu Gly
 283 450 455 460
 285 Pro Asn Ala Gly Glu Val Thr Gln Gly Phe Ala Leu Gly Ile Lys Cys
 286 465 470 475 480
 288 Gly Ala Ser Tyr Ala Gln Val Met Arg Thr Val Gly Ile His Pro Thr
 289 485 490 495
 291 Cys Ser Glu Glu Val Val Lys Leu Arg Ile Ser Lys Arg Ser Gly Leu
 292 500 505 510
 W-- 294 Asp Pro Thr Val Thr Gly Cys Xaa Gly
 295 515 520
 298 <210> SEQ ID NO: 3
 299 <211> LENGTH: 2056
 300 <212> TYPE: DNA
 301 <213> ORGANISM: Homo sapiens
 303 <220> FEATURE:
 304 <221> NAME/KEY: CDS
 305 <222> LOCATION: (188)..(1669)
 307 <220> FEATURE:
 308 <221> NAME/KEY: misc_feature
 309 <222> LOCATION: (1664)..(1666)
 310 <223> OTHER INFORMATION: "tga" is translated to selenocysteine
 312 <400> SEQUENCE: 3
 313 gtcccgacc tcaggccag ttcagtgtac ttccctctc tacttctcc ctccagtc 60
 315 ttctccatcc ctccctttt tggtgccccc ttgcctgcct tctcgccag tagcttg 120
 317 agtagacacg atgacacctt ttgcaggcta aaaaggctga gaggggcact atgtgcagt 180
 319 agccacc atg gag gac caa gca ggt cag cgg gac tat gat ctc ctg gtg 229
 320 Met Glu Asp Gln Ala Gly Gln Arg Asp Tyr Asp Leu Leu Val
 321 1 5 10
 323 gtc ggc ggg gga tct ggt ggc ctg gct tgt gcc aag gag gcc gcc cag 277
 324 Val Gly Gly Gly Ser Gly Gly Leu Ala Cys Ala Lys Glu Ala Ala Gln
 325 15 20 25 30
 327 ctg gga agg aag gtg gcc gtg gtg gac tac gtg gaa cct tct ccc caa 325
 328 Leu Gly Arg Lys Val Ala Val Val Asp Tyr Val Glu Pro Ser Pro Gln
 329 35 40 45
 331 ggc acc cgg tgg ggc ctc ggc ggc acc tgc gtc aac gtg ggc tgc atc 373
 332 Gly Thr Arg Trp Gly Leu Gly Gly Thr Cys Val Asn Val Gly Cys Ile
 333 50 55 60
 335 ccc aag aag ctg atg cac cag gcg gca ctg ctg gga ggc ctg atc caa 421
 336 Pro Lys Lys Leu Met His Gln Ala Ala Leu Leu Gly Gly Leu Ile Gln
 337 65 70 75
 339 gat gcc ccc aac tat ggc tgg gag gtg gcc cag ccc gtg ccg cat gac 469
 340 Asp Ala Pro Asn Tyr Gly Trp Glu Val Ala Gln Pro Val Pro His Asp
 341 80 85 90

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/830,706B

DATE: 04/22/2002
TIME: 15:01:39

Input Set : A:\55865sequence.txt
Output Set: N:\CRF3\04222002\I830706B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; Xaa Pos. 520
Seq#:2; Xaa Pos. 520
Seq#:3; Xaa Pos. 493
Seq#:4; Xaa Pos. 493
Seq#:37; N Pos. 31417
Seq#:38; Xaa Pos. 498

09830706-080704
10/2080-90/080860

VERIFICATION SUMMARY

DATE: 04/22/2002

PATENT APPLICATION: US/09/830,706B

TIME: 15:01:39

Input Set : A:\55865sequence.txt

Output Set: N:\CRF3\04222002\I830706B.raw

L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:169 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:1592
L:294 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:512
L:440 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:1669
L:559 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:480
L:2002 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:37 after pos.:31380
L:3281 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:38 after pos.:496

I830706B-080701